

# 1) Media Server

- **Erklären Sie den Begriff Quality of Service.**

Quality of Service (QoS) ist ein gerne verwendeter Begriff um die Programm-Erfordernisse bzw. Voraussetzungen für eine vorgegebene Ressource zu repräsentieren. Formal ist QoS eine Menge von Qualitätsanforderungen an das gemeinsame Verhalten bzw. Zusammenspiel von mehreren Objekten.

- **Erklären Sie kurz 3 Quality of Service Parameter für Media Server.**

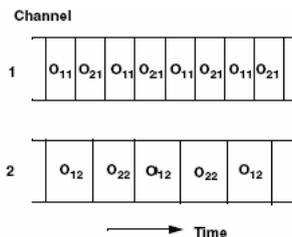
- Bandbreitenanforderungen: Die Zustellung von spezifischen Datenraten (muss nicht konstant sein, durchschnittliche Datenrate)
- Peak-Bandbreitenanforderungen: Maximale Dauer der Spitzenbandbreite
- Burst size oder Workahead: maximale Datenmenge, bei der der Stream eine strikt konstante Rate übersteigen kann
- Verzögerung: Maximale Verzögerung
- Verlustwahrscheinlichkeit: Maximale tolerierbare Verlustwahrscheinlichkeit (Kontrollinformation ist empfindlicher)

- **Was sind die Ziele des Channel Scheduling?**

- Minimierung der Absagewahrscheinlichkeiten auf längere Zeit
- Minimierung der Durchschnittswartezeit
- Unterschiedliche Requests fair behandeln („hot videos“ nicht bevorzugen)  
→ eine Schedulingstrategie ist dann fair, wenn die Absagewahrscheinlichkeit für alle Requests gleich ist
- Wartezeit bis zur Wiederaufnahme (zw. Aufnahmen und Abspielen)
- Berücksichtigung von VCR-Anforderungen
- → Erweiterung Contingency Channels, um Server Load zu minimieren

- **Beschreiben Sie kurz Pyramid Broadcasting. Wozu dient das Verfahren?**

Pyramid Broadcasting ist ein Verfahren zum Zusammenfassen mehrerer Anfragen in logische Kanäle, damit nicht für jede einzelne Anfrage ein neuer logischer Kanal geöffnet werden muss (→ wäre negativ für Performance). Während in einfachen NVOD (Near Video On Demand)-Strategien die Wartezeit durch die Anzahl\_der\_Objekte \* Abspielzeit\_der\_Objekte / Anzahl\_logischer\_Kanäle bestimmt ist, kann diese Zeit durch das Pyramid Broadcasting-Verfahren reduziert werden.



- Die Serverbandbreite wird in  $N_{pyramid.channel}$  logische Channels unterteilt; (Beispiel: 2 logische Channels, jeder mit 3 mal so viel Bandbreite, die für Wiedergabe gebraucht wird)

- Jedes Video wird in  $N_{pyramid.channel}$  Segmente geteilt, dessen Größe durch  $\alpha_{pyramid.channel}$  vergrößert wird.
- Wenn  $N_{pyramid.channel}$  größer wird, nimmt die Größe des 1. Segments ab
- Das i-te Segment aller Videos wird wiederholt Broadcast auf Channel i gesendet (= Unterschied zum simplen NVO: Videos werden über die Kanäle hinweg unterteilt, je ein Teil wird auf einem anderen Kanal gesendet)
- Um das Video i zu spielen, startet der Client den Download und spielt  $O_{i1}$  wenn es verfügbar ist.
- Das Schema benötigt ausreichend lokalen Speicher (dh. beim Player) in der Größenordnung von dem größten Segment
- Der Wert von  $\alpha_{pyramid.channel}$  : für ununterbrochene Anzeige, muss es möglich sein den Download des nächsten Segments zu starten, bevor das momentane Segment vollständig angezeigt wurde.
- Der maximale Wert von  $\alpha_{pyramid.channel} : \frac{N_{logical.server.chan}}{N_{Objects} * N_{pyramid.channel}}$

- **Beschreiben Sie kurz die Everest Contiguous Allocation. Wozu dient sie?**

Ziel ist, die Dateien auf der Platte so gut wie möglich in aufeinanderfolgende Blöcke aufzuteilen (und eine Fragmentierung zu vermeiden, da eine Defragmentierung sehr aufwendig ist).

Für einzelne Dateien können dabei Blöcke in Größen von  $w^i$  zugewiesen werden.  $w$  &  $i$  sind dabei zu definierende Parameter.

Beispiel: eine Datei mit einer Länge von 5 Blöcken kann folgendermaßen angelegt werden:

$w=2$ : 1 Segment mit 4 Blöcken ( $2^2$ ) und 1 Segment mit 1 Block

$w=3$ : 1 Segment mit 3 Blöcken und 2 Segmente zu je einem Block

- Freie Blöcke werden zu freien Segmenten zusammengefasst und in einer eigenen Liste geführt - pro Größe  $w$  jeweils eine Liste

- Maximal  $w-1$  Segmente in einer Liste

- Sollte es in Folge einer Löschung in einer Liste zu mehr als  $w-1$  Segmenten kommen, werden die Segmente so verteilt, dass zumindest  $w$  freie Segmente der Größe  $w^i$  aufeinanderfolgen.

Diese werden dann zusammengeführt in ein Segment der Größe  $w^{i+1}$

- **Welche Constrained-Placement Policies kennen Sie? Diskutieren kurz ein Verfahren.**

REBECA (Region based block allocation method)

Die Platten werden in eine fixe Anzahl von aufeinanderfolgenden Bereichen unterteilt. Alle Blöcke in einer Region werden mit einem Scan gelesen. Je größer die Anzahl der Regionen desto weniger Such-Overhead und desto höhere Beginn-Zugriffszeiten sind möglich.

Aufeinanderfolgende Blöcke eines Medienobjekts werden entsprechend der Reihenfolge der Bewegungen auf der Platte gespeichert.

### Strand-based allocation

Als Alternative zum MPEG-1/2-Ansatz ist es möglich, Audio- und Video-Datenströme getrennt voneinander speichern. Dabei ist es erforderlich die Daten nahe beieinander zu speichern, damit eine kontinuierliche Versorgung mit Daten gewährleistet werden kann. Voraussetzung ist, dass in strand\_1 entsprechend große Gaps für Media Blöcke des strand\_2 bereitstehen, sowie die Kontinuitätsbedingungen nicht verletzt werden.

→ Zugriffszeit & Datenübertragungsleistung muss ausreichend sein, um Unterbrechung durch die Blöcke von s1 zu kompensieren. Merge Condition  $G1/M2 \geq M1/G2$ .

Sonderbehandlung erforderlich, falls die beiden strand-patterns nicht genau zusammenpassen.

- **Beschreiben Sie kurz Group Sweeping Scheduling. Wozu dient das Verfahren?**

GSS ist eine Kombination der Vorteile Round Robin und Scan Methode:

1. Vorteil der Scan-Methode Seek-Overhead wird minimiert (was bei Round Robin nur durch zusätzlichen Buffer erreicht werden kann)
2. Vorteil der Round-Robin-Methode: benötigter Read-Ahead-Buffer (bei Scan) kann durch Aufteilung auf verschiedene Gruppen minimiert werden.

Unter starker Belastung tendiert GSS-Scheduling dazu genauso wie die Scan-Methode zu arbeiten. (S.97-100 in den Folien & S.163-165 Chapter 8 in den Unterlagen)

Grundsätzliches Ziel der drei Methoden ist, die zu lesenden Daten ohne Verzögerung einlesen zu können - das geschieht, indem die einzelnen Datenblöcke bereits im voraus gelesen werden, bevor sie überhaupt benötigt werden.

- **Beschreiben Sie kurz die Bandwidth-to-Space (BSR) Policy. Wozu dient das Verfahren?**

- Ziel ist, die vorhandenen Kapazitäten (Platz/Bandbreite) optimal durch entsprechende Verteilung der Medienobjekte auszunutzen
- Heiße und kalte sowie kleine und große Video-Objekte werden auf den Datenträgern zusammengefasst (striping groups)
- Die Datenträger werden dabei entsprechend dem Bandbreiten-Platz-Verhältnis betrachtet
- Analog dazu werden Medienobjekte klassifiziert
- BSR erstellt und löscht Kopien (Repliken) entsprechend Bedarf.
- Beispiel für schlechte Aufteilung: mehrere Objekte mit hohem Platzbedarf, aber geringer Bandbreite
- Beispiel für gute Aufteilung: mehrere Objekte in Summe entsprechend der Charakteristik des jeweiligen Geräts angepasst.

- **Beschreiben Sie kurz das General Interval Caching Policy. Worin unterscheidet sie sich vom Interval Caching?**

#### IC/GIC-Policy:

Datenblöcke, die sich zwischen einem Paar von aufeinanderfolgenden Streams (gebildet aus preceding und following Streams) befinden, werden als Intervall bezeichnet. Der nachfolgende Stream wird bei dem Verfahren durch die bereits benutzten Blöcke des Vorgängers versorgt.

Die Intervallgröße: Zeitdifferenz zwischen Zugriff der beiden Datenströme auf den gleichen Block. Nur die kürzesten Intervalle werden gecached (Max. cache hit rate & I/O Minimierung). Geringer Implementations-Overhead, Änderungen ergeben sich nur bei Ankunft oder Beendigung eines Stroms.

#### GIC-Besonderheiten:

IC eignet sich nicht für kurze Medienobjekte, wie z.B. Videoclips, da diese zu kurz sind, um Intervalle zu bilden. GIC erweitert deshalb die IC-Policy entsprechend für kleine Objekte: Die Intervalldefinition wird dahingehend erweitert, dass angenommen wird, dass die Objekte (für die Intervallberechnung) größer als tatsächlich sind. Unter dieser Annahme wird die Position des Vorgängerstreams ermittelt. Die Intervallgröße wird wie bei IC bestimmt, wobei zu beachten ist, dass die Cache-Erfordernisse geringer sind, denn der Clip ist ja kleiner als das ermittelte Intervall. Rest wie IC.

## 2) Visual Retrieval

- **Was ist die Brodatz-Datenbank und weshalb ist sie von Bedeutung?**

Die Brodatz-Datenbank ist der defacto-Standard für die Evaluierung von Textur-Algorithmen und besteht aus einer Datenbank mit 112 Graustufenbildern (512x512 8-bit), die jeweils eine unterschiedliche Texturklasse repräsentieren. Sie gibt an wie gut unterschiedliche Algorithmen Texturen in Bildern erkennen.

- **Beschreiben Sie kurz die Evaluierung von Texture-Features mit Hilfe der Brodatz-Datenbank.**

Aus der Mitte von jedem der 112 Brodatz-Bilder werden 9 Subbilder (128x128) gewonnen. Die Features werden berechnet, indem man den Textur-Algorithmus an jedem der insgesamt 1008 Subbilder anwendet. Danach wird die Distanz jedes Test-Subbildes zu allen anderen berechnet. Man erhält 2 Ergebnisse:

- Retrieval Rate:

Prozentanteil der n nächsten Subbilder, die zur selben Gruppe wie das Test-Subbild gehören.

Bsp.: Werden 10 Bilder gefunden, aber nur 7 dieser Bilder entsprechen einer bestimmten Klasse → Retrieval Rate = 70%

- Average Retrieval Rate:

Wird berechnet, indem jedes Subbild einmal als Test-Subbild errechnet wird. (d.h. alle Subbilder werden miteinander verglichen.)

- **Nennen Sie ein von geometrischen Transformationen unabhängiges Image-Feature. Diskutieren Sie kurz Vor- und Nachteile.**

Farbe bzw. die Verteilung von Farbe innerhalb von Bildern ist ein von geometrischen Transformationen unabhängiges Image-Feature. Wird in Form von Histogrammen dargestellt. Ein Histogramm gibt die Anzahl der Pixel je vorkommender Farbe an. Vorteil der Histogramme ist die Unabhängigkeit gegenüber Translation und Rotation entlang der View-Achse.

Größter Nachteil ist, dass ein Histogramm keine Aussage über Bildbereiche zulässt.

Lösung: → Aufteilung in Bildbereiche und Bildung von Histogrammen für diese Bereiche.

- **Was versteht man unter Motion Continuity und wie wird sie berechnet?**

Bewegung spielt in VideoMaterial eine große Rolle. Motion Continuity beschreibt, wie „glatt“ eine Bewegung abläuft in dem es das Verhältnis von Bewegung zu Geschwindigkeit berechnet.

Die Bilder werden in kleine Blöcke unterteilt. Die Bewegungsvektoren werden mittels block matching berechnet. Es wird also berechnet, wie sich die einzelnen Blöcke von Frame zu Frame verschoben haben. Verschieben sich die Blöcke einer einzelnen Region → Objekt bewegt sich; verschiebt sich alles (einheitlich) → Kamera bewegt sich.

- **Erklären Sie kurz die Funktionsweise des Twin-Comparison Approach.**

Twin-Comparison dient zur Erkennung von kontinuierlichen Übergängen. Während Breaks anhand einer Histogrammdifferenz von  $> T_B$  leicht erkannt werden kann, ist es zur Erkennung von kontinuierlichen Übergängen erforderlich, zusätzlich eine Grenze

$T_S$  einzuführen. Liegt die Histogrammdifferenz über  $T_S$ , wird der betroffene Frame als potentieller Start eines Übergangs markiert und die folgenden Differenzen so lange akkumuliert, bis die Differenz den Schwellwert  $T_B$  überschreitet. In dem Fall wurde ein Übergang erkannt und der Endframe wird als Ende der Überganssequenz markiert. Fällt die Differenz jedoch wieder unter  $T_S$  bevor  $T_B$  erreicht werden konnte, handelt es sich um keinen Übergang.

TCA benötigt die Verwendung von 2 Beschränkungs-Schwellwerten:  $T_B$  wird für die Erkennung eines Schnitts, und  $T_S$  für die Erkennung eines Special-Effekts verwendet. Wann immer eine Differenz zwischen  $T_B$  und  $T_S$  fällt, wird der entsprechende Frame markiert und mit den nachfolgenden Frames verglichen. Der End-Frame eines Übergangs wird entdeckt, wenn die Differenzen zwischen aufeinanderfolgenden Frames unter den Schwellwert von  $T_S$  sinkt, während der Vergleich zwischen dem markierten Frame über den Wert von  $T_B$  steigt.  $SD_{pq}$  = die Differenz zwischen aufeinanderfolgenden Frames, die durch das Differenz-Maß definiert wird.

$SD'_{pq}$  : berechnet, wenn  $T_B > SD_{pq} > T_S$ , die angehäuften Differenzen zwischen dem aktuellen Frame und dem potentiellen Start-Frame  $F_S$  eines Übergangs (Transition).

$T_S$  = der Schwellenwert um den Startframe eines Übergangs zu entdecken.

$T_B$  = der Schwellenwert um den Endframe  $F_E$  und Schnitte zu entdecken.

- **Beschreiben Sie kurz den Multi-Pass-Ansatz zur Videoschnitterkennung. Weshalb verwendet man diesen Ansatz?**

Man möchte damit die Bearbeitungszeit reduzieren (eine Art Filter, welche Frames in Ordnung sind und welche weiter untersucht werden müssen)

1. Durchgang: Tb wird niedriger angesetzt → Auflösung wird vorübergehend heruntersgesetzt (weniger frames /sek.) → großer Sprungfaktor: steigert SD während schrittweisen Übergängen → Sowohl Schnitte als auch Übergänge werden erkannt (manchmal auch zu viele)

Weitere Durchgänge: gesteigerte temporäre Auflösung wird verwendet um potentielle Segmentgrenzen zu erkennen

- **Erklären Sie kurz in welchem Kontext die Erkennung von Kamerabewegungen eine Rolle spielt. Erklären Sie kurz die prinzipielle Arbeitsweise der Erkennung.**

Veränderungen, die mit Spezialeffect-Übergängen assoziiert werden, müssen von Veränderungen durch Kamerabewegungen, unterschieden werden.

Veränderungen auf Grund von Kamerabewegungen tendieren dazu aufeinanderfolgende Differenzwerte von derselben Reihenfolge, wie die von schrittweisen Übergängen, hervorzurufen. Daher ist es notwendig die Muster von Bildbewegungen, die durch Kamerabewegung hervorgerufen werden, zu entdecken.

Es gibt zwei Zugänge:

- Bewegungsvektor-Analyse
- (Video X-Ray), Visualisierung des 1. Verfahrens

Motion Vector Analysis:

Die einzelnen Frames werden in Blöcke unterteilt, wobei jedem Block ein Bewegungsvektor zugewiesen wird. Daraus entsteht ein Vektorfeld, mit Hilfe dessen erkannt werden kann, ob es sich um

- panning oder tilting:  
Alle Vektoren zeigen in Richtung des Modalvektors (mit geringen Abweichungen)
- zooming (in oder out):  
Summe der Vektoren ist 0. Zoom in: Alle Vektoren zeigen vom Zentrum nach außen. Zoom out: umgekehrt. Spezialfall, wenn Zoom-Mittelpunkt außerhalb des Bild-Zentrums liegt. → mit Hilfe der Annahme, dass der Zoom-MP innerhalb des Bildes liegen muss gelöst
- um vordefinierte Bewegungsmuster handelt  
z.B. Bewegungen von Menschen, Tieren, Objekten. Dazu kann eine entsprechende Beschreibung verwendet werden.

Video X-Ray:

- Annahme, dass der Hintergrund entsprechende Unterscheidungsmerkmale bietet.
- Panning wird durch schräge Linien in der oberen Ansicht erkannt (unter Vorstellung eines Würfels).
- Tilting wird durch schräge Linien in der seitlichen Ansicht erkannt.
- Zooming wird durch auseinander- oder zusammenlaufende Linien erkannt.

- **Erklären Sie kurz 2 Ansätze zur Erkennung der Kameraoption „Zoom“ und nennen Sie Vor- und Nachteile der Ansätze.**
  - Der Kamerafokus ist im Zentrum des Frames → alle Bewegungsvektoren zeigen entweder zu diesem Zentrum (zoom out) oder gehen von diesem Zentrum nach außen (zoom in) → Vektoren heben einander auf  
Vorteil: Zoom leicht zu erkennen → Summe der Vektoren = 0  
Nachteil: Fokus muss im Frame-Zentrum liegen
  - Der Fokus liegt wo anders (aber innerhalb des Frames) → der Bewegungsvektor in der obersten Reihe in der linkensten Spalte & der in der untersten Reihe und der rechtensten Spalte sind invers zueinander → Zoom-Operation.  
Vorteil: flexibler bezüglich Fokus als 1. Ansatz
- **Beschreiben Sie kurz 3 Werkzeuge für die Inhaltsanzeige (Content Indication) von Video.**  
Content Indication ist der Prozess von explizitem Präsentieren eines Teils des Inhalts für ein besseres Verständnis.

#### **Video Icons, es gibt:**

- Icon, basierend auf einem Frame, der von einem Shot extrahiert wird, mit Pseudotiefe für die Repräsentation von Dauer, Richtungspfeilen und Zeichen für die Repräsentation von Objekt und Kamerabewegung.
- Entstehung eines Bildes, das den globalen visuellen Inhalt des Shots repräsentiert
- Herausragende Standbilder
- Videospace icon (3D Video Icons)
- Videomap

#### **Video map**

Shot-Referenz, Intensitäten-Histogramm, Intensitäts-Durchschnitt, Histogramm-Differenzen, Video X-Ray (Ansicht von oben), Video X-Ray (Ansicht von der Seite), Farb-Histogramm

#### **Keyframe-Extrahierung**

Vorteil: einfache Berechnung: wenn signifikante Inhaltsänderungen zwischen dem aktuellen Frame des Shots und dem letzten Keyframe auftreten, wird der aktuelle Frame als Keyframe ausgewählt – signifikante Merkmale sind Farbe, Textur und Bewegung

Nachteil: nicht sehr repräsentativ

- **Welche Verfahren zur Schnitterkennung in MPEG-komprimierten Videos kennen Sie? Erklären Sie kurz die Arbeitsweise eines dieser Verfahren.**

Algorithmen, die auf DCT-Koeffizienten basieren:

*Paarweise DCT Block Vergleiche*

Vorteil: da nur ein kleiner Teil der Video Frames I-Frames sind, wird die Berechnungszeit reduziert.

Nachteil: der Verlust von zeitlicher Auflösung kann zu falschen Positiven führen

Algorithmen, die auf Bewegungsvektoren basieren (250):

*Gemischter Ansatz zur Partitionierung (251):*

1.Schritt: DCT-Vergleich der I-Frames mit einem großen Skip-Faktor um Regionen von potentiellen Schnitten, Übergängen, Kameraoperationen oder Objektbewegungen zu erkennen

2.Schritt: DCT-Vergleich mit geringerem Skip-Faktor in der Umgebung von potentiellen Schnitten, Übergängen,...

→ eliminiert zuerst falsch erkannte

weitere Schritte: Bewegungsbasierte Vergleiche entweder auf dem gesamten Video oder auf den zuvor erkannten Sequenzen verifizieren die DCT-Resultate und verbessern die Genauigkeit

Evaluierung: MotionVektor und hybrides Verfahren besser um Schnitte zu erkennen (DCT hat auch falsche Schnitte erkannt – da nur I-Frames verwendet wurden); bei Übergängen hat die MotionVektor-Methode allerdings versagt.

- **Worin besteht Ihrer Meinung nach das Hauptproblem bei der Videosegmentierung?**

Segmentierung ohne Berücksichtigung des Inhalts. Algorithmen bieten nicht die Möglichkeiten einer manuellen Segmentierung durch einen Menschen → Qualität der Ergebnisse.

- **Was sind die Aufgaben von Block Placement Strategien?**

Der Seek Overhead ist der Hauptoverhead beim Retrieval von der Disk. Die Block Placement Policies versuchen diesen zu reduzieren.

- **Wozu dienen Motionfeatures? Beschreiben Sie kurz 2 dieser Features.**

???

- **Erklären Sie kurz die Ziele von MPEG-7.**

- Standardisierte Beschreibung von Multimedia-Inhalten
- Flexibilität im Datenmanagement
- Globalisierung und Interoperabilität von Datenressourcen
- Standardisierung
  - Eines Sets von Description Schemes und Descriptors um Daten zu beschreiben
  - Einer Sprache um Description Schemes zu spezifizieren wie die Description Definition Language (DDL)
  - Eines Schemas zur Kodierung der Description

- **Erklären Sie kurz die Konzepte Deskriptor, Description Scheme und DDL im Kontext von MPEG-7.**

Deskriptor (D): Beschreiben die Merkmale (Features) von audiovisuellen Daten semantisch und syntaktisch. Definiert z.B.: Datentyp, erlaubte Werte (Wertebereich)  
Mehrere verschiedene Deskriptoren können ein einzelnes Feature darstellen.

Description Scheme (DS): Weist den Deskriptoren einen Kontext bzw. eine Bedeutung zu. Definiert die Struktur und Semantik von Beziehungen zwischen den Komponenten. Die Komponenten des Description Schemes können dabei sowohl Deskriptoren als auch Description Schemes sein. Im Gegensatz zum Deskriptor definiert das Description Scheme die Struktur einer Description, während ein Deskriptor ein Feature beschreibt.

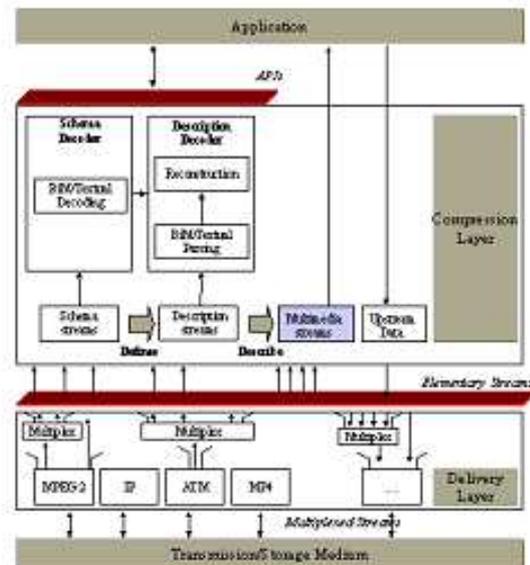
Description Definition Language (DDL): Beschreibt die Syntax der Description Schemes.  
Erlaubt das Erstellen von neuen Description Schemes und Deskriptoren. Auch (applikationsspezifische) Erweiterungen, Veränderungen und das Kombinieren vorhandener Description Schemes und Deskriptoren sind erlaubt. DDL basiert auf XML, ist aber keine Modellierungssprache!

- **Erklären Sie kurz die Terminal-Architektur von MPEG-7.**

Terminal ist das Teil, das die kodierten Repräsentationen des Multimedia-Contents benutzt. Dabei kann es sich um eine eigenständige Applikation oder einen Teil eines Applikationssystems handeln.

Die Architektur besteht aus:

- Application
- Compression Layer  
hier werden die Access-Units<sup>1</sup> gepackt und die content description wird rekonstruiert - eine Konvertierung ins Textformat im Rahmen der Decodierung ist nicht Voraussetzung. Verarbeitung kann entweder im Textformat oder in einem proprietären Binärformat erfolgen.
- Delivery Layer  
stellt Mechanismen für die Synchronisation, Framing und Multiplexing von MPEG-7 Inhalten bereit.
  - MPEG-7 inhalte können unabhängig oder mit dem Content selbst übertragen werden
  - Nicht alle MPEG-7 Streams müssen downstream-fähig sein.
  - Stellt elementary Streams für den Compression Layer zur Verfügung.
  - Elementary streams bestehen aus aufeinanderfolgenden einzeln zugreifbaren Datenpaketen → Access Units (= Kleinste Dateneinheit, der Zeitinformationen zugewiesen werden können)
  - Elementary Streams beinhalten folgende Informationen
    - Schema information - Struktur der MPEG-7 Description
    - Descriptions information - komplette oder teilweise Beschreibung des Contents
- Transmission / Storage Medium  
unterste Ebene, liefert gemultiplexte Streams an den Delivery Layer (DL)



<sup>1</sup> Kleinste Data-Entity in der in der die timing information hinzugefügt werden kann

- **Erklären Sie kurz die Konzepte Rights Data Dictionary (RDD) und Rights Expression Language (REL). In welchem Kontext treten sie auf und wie spielen sie zusammen?**

Die RDD und REL treten im Kontext von MPEG-21 auf. REL benutzt das Rights Data Dictionary (RDD) als Vokabular. Die Rights Expression Language ist ein XML-Schema (daher maschinenlesbar!) zur Definition von verschiedenen Benutzerrechten für ein Digital Item<sup>2</sup>.

Der Schutz von Rechten insbesondere des Urheberrechts ist ein wesentlicher Bestandteil von MPEG-21. Die REL soll daher digitale Inhalte schützen, beabsichtigt die Spezifikation von Kontrolle und der Benutzung von digitalen Inhalten und soll Bedingungen für den Austausch von sensiblen oder privaten digitalen Inhalten unterstützen.

REL kennt folgende Rechte:

- Abspielen
- Ansehen
- Ausdrucken
- Editieren

Außerdem kennt sie verschiedene Conditions:

- Gültigkeit innerhalb bestimmter Zeiträume
- Bezahlarten wie pay-per-use oder flat-fee
- Nutzungsanzahlen wie einmalige oder unbegrenzte Nutzung
- Gebietsgültigkeiten wie Europa oder Amerika

Das Rights Data Dictionary (RDD) definiert verschiedene Bedingungen aus dem Bereich digitaler Rechteverwaltung. Die hier eingetragenen Begriffe sollen eindeutig sein und sind mit einer Auslegung versehen, damit es keine Probleme bei der Implementierung gibt.

- **Erklären Sie kurz die Ziele von MPEG-21?**

- „to define a multimedia framework to enable transparent and augmented use of multimedia resources across a wide range of networks and devices used by different communities.“ (Vision der Arbeitsgruppe des MPEG-21 Standards)
- Einheitliches Framework, das alle Aspekte der Erstellung, Verteilung und Konsumierung von Multimediainhalten regelt.
- Schaffung einer system- und plattformunabhängigen Umgebung für den Austausch und Gebrauch von Medieninhalten.
- Soll die Integration von multimedialen Inhalten verschiedener Formate (Standards) ermöglichen.
- Vereinheitlichung des Multimedieverkehrs: Multimedia Content Provider mit verschiedenen Geschäftsmodellen sollen auf eine einheitliche Infrastruktur von der Erstellung bis zur Konsumierung von Multimedialen Inhalten zurückgreifen können.
- Effiziente und automatisierte Interoperabilität zwischen verschiedenen Multimedia-Dienstleistern.
- Standardisierung sämtlicher Komponenten

---

<sup>2</sup> ist ein strukturiertes, digitales Objekt mit einer standardisierten Repräsentation, Identifikation und Metadaten innerhalb des MPEG-21 Standards; sie sind fundamentale Verarbeitungseinheiten (Transaktion, Verteilung) innerhalb des Frameworks; werden durch Digital Item Declaration Language (DIDL) beschrieben (XML!)

- Soll Benutzern, Herstellern von Multimediainhalten und -diensten Homogenität, Transparenz und Kompatibilität bieten.
  - Integration bereits bestehender jedoch komplexer Technologien zu einem Gesamtsystem (z.B.: Digital Rights Management (DRM))
- **Erklären Sie kurz das Konzept Segment Description Scheme und seine Anwendung.**  
 Das Segment DS ist das core element der Content Description. Mit dem Segment DS werden zeitliche und räumliche Aspekte beschrieben. Es adressiert die Beschreibung von physikalischen und logischen Aspekten von audiovisuellem Inhalt. Das Segment DS wird verwendet um segment trees zu erstellen. Das Segment DS ist eine abstrakte Klasse und hat 5 Unterklassen: AudioVisual Segment DS, Audio Segment DS, Still Region DS, Moving Region DS and Video Segment DS. (*Anmerkung: 5 Unterklassen lt. Folien; lt. anderen Quellen sind es 9!*)  
 Bsp.: Ein zeitliches Segment kann eine Menge von Samples sein, die von dem Audio Segment DS beschrieben wird. Ein räumliches Segment kann ein Teil eines Bildes oder Frames in einer Videosequenz sein, das durch das Still Region DS beschrieben wird.  
 Pro Segment können CreationInformation, Verwendung, Media Location und Text Anmerkungen angegeben werden.
  - **Was versteht man unter Variation Description Schemes und wofür werden sie verwendet?**  
 Ermöglicht Servern eine Alternative basierend auf Möglichkeiten des Terminals, Netzwerkmodalitäten und Präferenzen der Benutzer auszuwählen. Das Variation DS spezifiziert die verschiedenen Variationen von audiovisuellem Inhalt (AV-Inhalt) wie Zusammenfassungen, komprimierte oder niedrig aufgelöste Version, verschiedene Sprachen und Modalitäten (Audio, Video, Bild, Text). Eine der Hauptfunktionen vom Variation DS ist dem Server oder Proxy eine passende Variation des audiovisuellen Inhalts zu den Fähigkeiten der Endgeräte, Netzwerkbedingungen und Benutzervorlieben zu liefern. Genau diese Variationen werden vom Variation DS beschrieben. Die Variationen beziehen sich auf erst kürzlich gelieferten AV-Inhalt oder auf AV-Inhalt anderer Quellen. Die variation fidelity value gibt die Qualität im Vergleich zum Original an. Das variation type attribute gibt die Art (Zusammenfassung, Auszug, Sprachübersetzung, Farbreduktion, Kompression...) der Veränderung an.
  - **Navigation in MPEG-7 (*genauer Wortlaut der Frage unbekannt*)**  
 MPEG-7 stellt DSs zu Verfügung, die Navigation und Zugang von audiovisuellem Inhalt unterstützen. Folgendes wird spezifiziert:
    - Summaries – aktiviert effizientes Browsen, navigieren, entdecken, visualisieren...
    - Views and Partitions – erlaubt multi-resolution und progressive access
    - Variations – spezifiziert die Relationen zwischen verschiedenen AV-Inhalten, die eine Auswahl verschiedener Variationen von Inhalt unter verschiedenen Lieferbedingungen erlaubt.

## Quellen:

- Folien vom Vortragenden SS 2007
- Folien von der Vorlesung „Multimediale Datenbanken“ WS06/07 (MMDB-VO05-06.pdf)
- Prüfungsordner von MM1: MM1\_VO\_TU\_Breiteneder\_PO\_2003-2005.pdf
- Prüfungsordner von MM2: gesamt.doc, mm2\_technologien\_tu\_breiteneder\_ausarbeitung\_2006-10-01.pdf, multimedia2\_technologie-po\_zusammenfassung.pdf, MM2-PO-Ausarbeitung\_d5.doc, MM2-PO-Ausarbeitung.doc
- <http://de.wikipedia.org/wiki/MPEG-21>
- [http://en.wikipedia.org/wiki/Description\\_Definition\\_Language](http://en.wikipedia.org/wiki/Description_Definition_Language)
- [http://de.wikipedia.org/wiki/Quality\\_of\\_Service](http://de.wikipedia.org/wiki/Quality_of_Service)
- <http://www.gi-ev.de/service/informatiklexikon/informatiklexikon-detailansicht/meldung/52/>
- <http://www.semanticmedia-showcase.de/WerkstattCM/CrossMedia/Ausarbeitungen/MPEG-21%20Ausarbeitung.pdf>

## Hinweise:

Dieser PO ist mehr oder weniger nur eine Zusammenfassung verschiedener bestehender Prüfungsordner bzw. Prüfungsfragen und versteht sich als Erweiterung bzw. Ergänzung zum PO von Waxy -> <http://www.informatik-forum.at/attachment.php?attachmentid=8827&d=1159713759>

Zum neuen Thema Audio Retrieval gibt es noch keine Fragen. Weiters sind wahrscheinlich noch nicht alle Fragen zu MPEG-7, DVB und MHP enthalten.

Keine Gewähr auf Richtigkeit oder Vollständigkeit.

Letzte Änderung: 21.10.2007